

Use of this site is subject to express terms of use. By continuing past this page, you agree to abide by these terms.

[Give us feedback](#) - Discuss with others on [google.public.labs.glossary](http://google.public.labs.glossary)



Look up **urn** at [Dictionary.com](#) or [Merriam-Webster](#)

Related phrases

[URN / Uniform Resource Name](#)  
[Uniform Resource Name \(URN\)](#)

[Bucket urn](#)  
[Classical urn](#)

[Globular urn](#)  
[Barrel urn](#)

[Cinerary urn](#)  
[cemetery urn](#)

Definitions for **URN** from the web

- A pottery vessel used for domestic, storage or burial purposes. Several distinct types are known: Collared urn Biconical urn Encrusted urn Bucket urn See Grooved ware, Peterborough ware  
[http://www.halifax-today.co.uk/specialfeatures/triviatrail/a22\\_u.html](http://www.halifax-today.co.uk/specialfeatures/triviatrail/a22_u.html)
- Uniform Resource Number. These will eventually replace URLs. A URN will automatically find the file you want from the closest or least busy server where the file is mirrored/cached. You can think of a URN as analogous to extending the ISBN book cataloging system to cover all the web pages as well. If you think about how paper books are distributed through wholesalers, bookstores and libraries, that will give you a metaphorical framework with which to think about URN caching. It works on much the same economic principles. The main difference is that with a URN, it is very cheap to Xerox a whole "book". Whoever is in charge of URNs have been dragging their feet for years. People are getting impatient waiting. Napster is a crude attempt to do this for music files.  
<http://209.139.205.39/jglossu.html>
- Uniform Resource Name, the result of an evolving attempt to define a name and address syntax for persistent objects accessible over the Internet; urn:foo:a123,456 is a legal URN consisting of three colon-separated fields: urn followed by a namespace identifier, followed by a namespace specifier (see RFC 1737 and RFC 2141 for details).  
<http://www.docbook.org/tdg/en/html/dbgloss.html>
- Uniform Resource Name. A standardized name for a persistent, location-independent, resource identifier. As conceived, when the system is implemented, you will be able to link to a resource by URN without specifying its location.  
<http://www.wondersky.com/etech/miscellaneous/glossinternet/u.htm>
- Uniform Resource Name. A URN identifies a resource or unit of information. It may identify, for example, intellectual content, a particular presentation of intellectual content, or whatever a name assignment authority determines is a uniquely namable entity.  
[http://weblinkers.web.cern.ch/WebLinker/WebLinker\\_6.html](http://weblinkers.web.cern.ch/WebLinker/WebLinker_6.html)
- Uniform Resource Name. A unique identifier that identifies an entity, but doesn't tell where it is located. A system can use a URN to look up an entity locally before trying to find it on the Web. It also allows the Web location to change, while still allowing the entity to be found.  
<http://www.wondersky.com/etech/miscellaneous/glossjava/u.htm>
- Uniform Resource Name - mechanism for uniquely identifying a resource "It is likely that in the future a browser will request a URN rather than a URL" (B.Kelly)  
<http://educate.lib.chalmers.se/cth/pathciv/internet/gloss.html>

- Uniform Resource Name. The next generation of document addresses will be handled by a URN server that will maintain the current URL to an object. Users, however, will refer to objects by name.  
[http://archive.ncsa.uiuc.edu/SDG/Software/mosaic-w/releaseinfo/2.1/WBook\\_72.html](http://archive.ncsa.uiuc.edu/SDG/Software/mosaic-w/releaseinfo/2.1/WBook_72.html)
- Uniform Resource Names, are as yet defined, but are the holy grail of addressing, as any file would retain the same URN, regardless of which computer the file resided on. URNs would be universal identifiers for Internet resources, regardless of the resource origins.  
<http://www.wiley.com/legacy/compbooks/graham/gloss.html>

---

**NAME • SYNOPSIS • DESCRIPTION • COMMANDS • SCHEMES • EXTENDING • CREDITS**  
• **KEYWORDS**

---

## NAME

**uri** -  
URI utilities

## SYNOPSIS

```
package require Tcl 8.2

package require uri ?1.1.1?

uri::split  url
uri::join   ?key value...
uri::resolve base url
uri::isrelative url
uri::geturl  url ?options...
uri::canonicalize uri
uri::register schemeList script
```

## DESCRIPTION

This package contains two parts. First it provides regular expressions for a number of url/uri schemes. Second it provides a number of commands for manipulating urls/uris and fetching data specified by them. For the latter this package analyses the requested url/uri and then dispatches it to the appropriate package (http, ftp, ...) for actual fetching.

## COMMANDS

**uri::split** takes a single *url*, decodes it and then returns a list of key/value pairs suitable for **array set** containing the constituents of the *url*. If the scheme is missing from the url it defaults to **http**. Currently only the schemes **http**, **ftp**, **mailto**, **urn** and **file** are supported. See section **EXTENDING** on how to expand that range.

**uri::join** takes a list of key/value pairs (generated by **uri::split**, for example) and returns the canonical url they represent. Currently only the schemes **http**, **ftp**, **mailto**, **urn** and **file** are supported. See section **EXTENDING** on how to expand that range.

**uri::isrelative** determines whether the specified *url* is absolute or relative.

**uri::resolve** resolves the specified *url* relative to *base*. In other words: A non-relative *url* is returned unchanged, whereas for a relative *url* the missing parts are taken from *base* and prepended to it. The result of this operation is returned. For an empty *url* the result is *base*.

**uri::geturl** decodes the specified *url* and then dispatches the request to the package appropriate for the scheme found in the url. The command assumes that the package to handle the given scheme either has the same name as the scheme itself (including possible capitalization) followed by `::geturl`, or, in case of this failing, has the same name as the scheme itself (including possible capitalization). It further assumes that whatever package was loaded provides a `geturl`-command in the namespace of the same name as the package itself. This command is called with the given *url* and all given *options*. Currently `geturl` does not handle any options itself.

**Note:** file-urls are an exception to the rule described above. They are handled internally.

It is not possible to specify results of the command. They depend on the `geturl`-command for the scheme the request was dispatched to.

**uri::canonicalize** returns the canonical form of a URI. The canonical form of a URI is one where relative path specifications, ie. . and .., have been resolved.

**uri::register** registers the first element of *schemeList* as a new scheme and the remaining elements as aliases for this scheme. It creates the namespace for the scheme and executes the *script* in the new namespace. The script has to declare variables containing the regular expressions relevant to the scheme. At least the variable `schemepart` has to be declared as that one is used to extend the variables keeping track of the registered schemes.

## SCHEMES

In addition to the commands mentioned above this package provides regular expression to recognize urls for a number of url schemes.

For each supported scheme a namespace of the same name as the scheme itself is provided inside of the namespace `uri` containing the variable `url` whose contents are a regular expression to recognize urls of that scheme. Additional variables may contain regular expressions for parts of urls for that scheme.

The variable `uri::schemes` contains a list of all supported schemes. Currently these are `ftp`, file, `http`, `gopher`, `mailto`, `news`, `wais` and `prospero`.

## EXTENDING

Extending the range of schemes supported by **uri::split** and **uri::join** is easy because both commands do not handle the request by themselves but dispatch it to another command in the **uri** namespace using the scheme of the url as criterion.

**uri::split** and **uri::join** call `Split[string totitle <scheme>]` and `Join[string totitle <scheme>]` respectively.

## CREDITS

Original code by Andreas Kupries. Modularisation by Steve Ball.

## KEYWORDS

uri, url, fetching information, www, http, ftp, mailto, gopher, wais, prospero, file

---

[Table of Contents](#) • [Index](#) • [Keywords](#)

Use of this site is subject to express terms of use. By continuing past this page, you agree to abide by these terms.

[Give us feedback](#) - Discuss with others on [google.public.labs.glossary](http://google.public.labs.glossary)



Look up **uri** at [Dictionary.com](#) or [Merriam-Webster](#)

Related phrases

[namespace URI](#) [Url / Uri](#) [Base-URI](#) [Labelled URI](#) [Page Request\(s\)](#) [URI Too Large](#)

Definitions for **URI** from the web

- (Uniform Resource Identifier) The super-set of URNs, URLs and URCS. URI is the way you identify any of the components of web content, whether it be a page of text, a video or sound clip, a still or animated image, or a program. The most common form of URI is the Web page address, which is a particular form or subset of URI called a Uniform Resource Locator (URL).  
<http://www.education.tas.gov.au/wiseweb/appendices/glossary.htm>
- Uniform Resource Identifier. A compact string of characters for identifying an abstract or physical resource. A URI is either a URL or a URN. URLs and URNs are concrete entities that actually exist; A URI is an abstract superclass  
<http://www.wondersky.com/etech/miscellaneous/glossjava/u.htm>
- A "Universal Resource Identifier". A URI is either a URL or a URN. (URLs and URNs are concrete entities that actually exist. A "URI" is an abstract superclass -- it's a name we can use when we know we are dealing with either an URL or an URN, and we don't care which.  
<http://java.sun.com/xml/jaxp/dist/1.0.1/docs/tutorial/glossary.html>
- Universal Resource Identifier used to label Web objects. A URI doesn't imply anything about the properties of an object such as its name and/or address. URLs, URNs, and URCS are the different forms of a URI.  
[http://archive.ncsa.uiuc.edu/SDG/Software/mosaic-w/releaseinfo/2.1/WBook\\_72.html](http://archive.ncsa.uiuc.edu/SDG/Software/mosaic-w/releaseinfo/2.1/WBook_72.html)
- A "Universal Resource Identifier". A URI is either a URL or a URN. (URLs and URNs are concrete entities that actually exist. A "URI" is an abstract superclass -- it's a name we can use when we know we are dealing with either an URL or an URN, and we don't care which.  
<http://industry.ebi.ac.uk/~senger/xml/docs/tutorial/glossary.html>
- Uniform Resource Identifier, the W3C's codification of the name and address syntax of present and future objects on the Internet. In its most basic form, a URI consists of a scheme name (such as file, http, ftp, news, mailto, gopher) followed by a colon, followed by a path whose nature is determined by the scheme that precedes it (see RFC 1630). URI is the umbrella term for URNs, URLs, and all other Uniform Resource Identifiers.  
<http://www.docbook.org/tdg/en/html/dbgloss.html>
- Uniform Resource Identifier, a string of characters that represents the location or address of a resource on the Internet and how that resource should be accessed. A URI is a superset of the Uniform Resource Locator.  
<http://www.brokertech.net/glossary3.htm>
- Uniform Resource Identifier. URLs are a type of URI.  
<http://docs.rinet.ru:8083/WebLomaster/appa.htm>

- Uniform Resource Identifier, is the filename part of a URL. If the URL is <http://www.hans.org/products/index.html>, then the URI is products/index.html. URN, URL.  
<http://209.139.205.39/jglossu.html>

Use of this site is subject to express terms of use. By continuing past this page, you agree to abide by these terms.

[Give us feedback](#) - Discuss with others on [google.public.labs.glossary](http://google.public.labs.glossary)



Look up **parsing** at [Dictionary.com](#) or [Merriam-Webster](#)

#### Related phrases

[Left-to-right parsing](#)  
[parsing order](#)  
[LR Parsing](#)

[Parsing Paradox](#)  
[gene parsing](#)  
[perceptual parsing](#)

[Error parsing web](#)  
[Bottom Up Parsing](#)  
[LALR parsing](#)

[Syntax Directed Parsing](#)  
[Top Down Parsing](#)

#### Definitions for **Parsing** from the web

- Parsing may be divided into parts: lexical analysis and semantic parsing. Lexical analysis divides strings into components based on punctuation or tagging. Semantic parsing then attempts to determine the meaning of the string.  
<http://dublincore.sedic.es/glossary.shtml.htm>
- The term 'parsing' is a process of string manipulation or inspection to determine (or change) its contents.  
<http://www.logisol.com.au/mraddoc/gstarted/mmanp59.htm>
- Parsing is the act whereby a document is scanned, and the information contained within the document is filtered into the context of the elements in which the information is structured.  
<http://orworld.uni-paderborn.de/downloads/glossary/glossary.html>
- The process or result of making a syntactic analysis (\*) (+) parser: tool (often automatic or semi-automatic computer program) used for parsing. ( General Description of Parsers - external link) parsed corpus: a corpus that have been syntactically analysed and provided with annotation representing the analysis.  
[http://clwww.essex.ac.uk/w3c/corpus\\_ling/content/glossary.html](http://clwww.essex.ac.uk/w3c/corpus_ling/content/glossary.html)
- We say that certain Emacs commands parse words or expressions in the text being edited. Really, all they know how to do is find the other end of a word or expression. See section The Syntax Table.  
[http://doc.ctrlaltdel.ch/emacs/manuel/emacs\\_495.html](http://doc.ctrlaltdel.ch/emacs/manuel/emacs_495.html)
- We say that Emacs parses words or expressions in the text being edited. Really, all it knows how to do is find the other end of a word or expression. See section 27.5 The Syntax Table.  
[http://www.la.utexas.edu/lab/software/user/gnu/xemacs/xemacs\\_31.html](http://www.la.utexas.edu/lab/software/user/gnu/xemacs/xemacs_31.html)
- The process by which a program written in some programming language is broken down into its syntactic elements.  
<http://www.erc.msstate.edu/~cass/cost/doc/CPlusPlusGlossary.html>
- Separating data elements; a part of the data hygiene process. Example: placing first name and last name in separate fields.  
<http://www.unitedwire.net/buzzwords/buzzdf01.htm>
- The process of deciphering input data and formatting it so that it can be used in a program.  
<http://artcontext.org/edu/topics/jargon/glossary.html>